

# SEQUENCE LISTING

## (1) GENERAL INFORMATION:

- (i) APPLICANT: KABUSHIKI KAISHA HAYASHIBARA SEIBUTSU KAGAKU  
KENKYUJO
- (ii) TITLE OF INVENTION: POLYPEPTIDE HAVING  $\beta$ -FRUCTOFURANOSIDASE  
ACTIVITY
- (iii) NUMBER OF SEQUENCES: 6
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: BROWDY AND NEIMARK
  - (B) STREET: 419 Seventh Street, N.W., Suite 300
  - (C) CITY: Washington
  - (D) STATE: D.C.
  - (E) COUNTRY: USA
  - (F) ZIP: 20004
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Floppy disk
  - (B) COMPUTER: IBM PC compatible
  - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
  - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER: 08/870,827
  - (B) FILING DATE: 06-JUN-1997
  - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: JP 170,630/1996
  - (B) FILING DATE: 10-JUN-1996
- (viii) ATTORNEY/AGENT INFORMATION:
  - (A) NAME: BROWDY, Roger L.
  - (B) REGISTRATION NUMBER: 25,618
  - (C) REFERENCE/DOCKET NUMBER: TSUSAKI=2
- (ix) TELECOMMUNICATION INFORMATION:
  - (A) TELEPHONE: 202-628-5197
  - (B) TELEFAX: 202-737-3528

## (2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear
- (iii) Fragment type: N-terminal fragment
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Met	Asn	Ser	Gly	Asp	Tyr	Lys	Glu	Asp	Tyr	Gly	Phe	Ala	His	Ile	Thr
1				5					10					15	
Arg	Ala	Asp	Met	Leu											
			20												

## (2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 12 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

(iii) Fragment type: internal fragment

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Ser Val Phe Asp Gly Gly Asp Gly Thr Val Tyr Gln  
1 5 10

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 455 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Met Asn Ser Gly Asp Tyr Lys Glu Asp Tyr Gly Phe Ala His Ile Thr  
1 5 10 15  
Arg Ala Asp Met Leu Lys Ile Pro Gly Gln Gln Asn Ser Pro Gln Phe  
20 25 30  
Lys Val Pro Gln Phe Asn Ala Ser Ala Ile Lys Asn Ile Asp Ser Ala  
35 40 45  
Lys Gly Tyr Asp Lys Ser Gly Asn Leu Ile Asp Leu Asp Val Trp Asp  
50 55 60  
Ser Trp Pro Leu Gln Asn Ala Asp Gly Thr Ala Ala Asn Tyr His Gly  
65 70 75 80  
Tyr His Ile Val Ser Ala Leu Ala Gly Asp Pro Lys Asn Ser Asp Asp  
85 90 95  
Thr Pro Leu His Leu Phe Tyr Gln Lys Val Gly Asp Thr Ser Ile Asp  
100 105 110  
Ser Trp Lys Asn Ala Gly Arg Val Phe Glu Asp Met Asp Lys Phe Val  
115 120 125  
Pro Asn Asp Pro Tyr Leu Lys Tyr Gln Thr Gln Glu Trp Ser Gly Ser  
130 135 140  
Ala Thr Leu Thr Lys Asp Gly Gln Val Arg Leu Phe Tyr Thr Asp Tyr  
145 150 155 160  
Ser Gly Asn Pro Glu Asp Gly Gly Thr Gly Ala Gly Asn Gln Ile Ile  
165 170 175  
Ser Thr Ala Gln Val Asn Leu Ser Gln Pro Asp Ala Ala Thr Leu Lys  
180 185 190  
Val Asp Gly Val Ser Asp His Lys Ser Val Phe Asp Gly Gly Asp Gly  
195 200 205  
Thr Val Tyr Gln Asn Ile Gln Gln Phe Ile Asp Glu Gly Lys Trp Ile  
210 215 220  
Ser Gly Asp Asn His Thr Leu Arg Asp Pro His Tyr Val Glu Asp Lys  
225 230 235 240  
Gly His Lys Tyr Leu Val Phe Glu Ala Asn Thr Gly Thr Thr Asp Gly  
245 250 255  
Tyr Gln Gly Asp Gln Ser Phe Asn Asn Lys Ala Tyr Tyr Gly Gly Ser  
260 265 270  
Asp Val Phe Phe Gln Asn Glu Lys Asn Lys Leu Leu Gln Ser Pro Lys  
275 280 285  
Lys Gln Ile Ala Ser Leu Ala Asn Gly Ala Leu Gly Ile Val Glu Leu  
290 295 300  
Ala Asp Asp Tyr Thr Val Lys Ser Val Met Lys Pro Leu Val Ala Ser  
305 310 315 320  
Asn Thr Val Ala Asp Glu Val Glu Arg Ala Asn Ile Phe Lys Met Asn  
325 330 335  
Asn Lys Trp Tyr Leu Phe Thr Asp Ser Arg Gly Ser Lys Met Thr Ser  
340 345 350  
Asp Gly Ile Asn Asp Lys Asp Val Tyr Met Leu Gly Pro Gly Gly Asp  
355 360 365  
Ser Leu Asn Gly Pro His Asn Pro Ile Asn Glu Thr Gly Leu Val Leu  
370 375 380

Asn Met Asn Leu Asp Pro Ala Asp Leu Thr His Thr Tyr Ser His Cys  
 385 390 395 400  
 Gly Ile Pro His Pro Glu Gly Asn Asn Val Val Leu Thr Ser Tyr Met  
 405 410 415  
 Thr Asn Arg Gly Phe Tyr Pro Glu His His Ser His Leu Arg Asp Lys  
 420 425 430  
 Leu Gly Val Asn Ile Lys Gly Ser Asp Thr Ser Gly Gly Glu Asn Ser  
 435 440 445  
 Ser Gly Gln Gly Gln Phe Pro  
 450 455

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1365 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

ATG AAC AGC GGG GAC TAC AAG GAA GAC TAT GGT TTT GCC CAT ATT ACA	48
Met Asn Ser Gly Asp Tyr Lys Glu Asp Tyr Gly Phe Ala His Ile Thr	
1 5 10 15	
CGC GCT GAC ATG CTA AAA ATT CCA GGA CAA CAA AAC AGT CCT CAA TTT	96
Arg Ala Asp Met Leu Lys Ile Pro Gly Gln Gln Asn Ser Pro Gln Phe	
20 25 30	
AAA GTG CCT CAA TTC AAT GCA TCA GCA ATC AAA AAC ATT GAT TCG GCA	144
Lys Val Pro Gln Phe Asn Ala Ser Ala Ile Lys Asn Ile Asp Ser Ala	
35 40 45	
AAA GGG TAT GAT AAG TCA GGC AAC TTA ATA GAT TTA GAT GTA TGG GAT	192
Lys Gly Tyr Asp Lys Ser Gly Asn Leu Ile Asp Leu Asp Val Trp Asp	
50 55 60	
AGC TGG CCA CTG CAA AAC GCT GAT GGT ACT GCG GCA AAT TAT CAT GGA	240
Ser Trp Pro Leu Gln Asn Ala Asp Gly Thr Ala Ala Asn Tyr His Gly	
65 70 75 80	
TAT CAC ATC GTC TCC GCT TTA GCA GGT GAC CCA AAA AAC AGT GAT GAT	288
Tyr His Ile Val Ser Ala Leu Ala Gly Asp Pro Lys Asn Ser Asp Asp	
85 90 95	
ACT CCA CTT CAT TTA TTC TAT CAA AAA GTC GGT GAT ACA TCG ATT GAC	336
Thr Pro Leu His Leu Phe Tyr Gln Lys Val Gly Asp Thr Ser Ile Asp	
100 105 110 115	
AGC TGG AAA AAT GCT GGA AGA GTA TTT GAA GAT ATG GAT AAA TTT GTT	384
Ser Trp Lys Asn Ala Gly Arg Val Phe Glu Asp Met Asp Lys Phe Val	
120 125 130	
CCA AAT GAT CCG TAT CTT AAA TAT CAA ACA CAG GAG TGG TCA GGT TCT	432
Pro Asn Asp Pro Tyr Leu Lys Tyr Gln Thr Gln Glu Trp Ser Gly Ser	
135 140 145	
GCT ACT TTA ACC AAA GAT GGC CAA GTC CGT TTA TTC TAT ACA GAT TAC	480
Ala Thr Leu Thr Lys Asp Gly Gln Val Arg Leu Phe Tyr Thr Asp Tyr	
150 155 160	
TCA GGT AAT CCT GAA GAT GGT GGA ACC GGT GCT GGT AAC CAA ATC ATT	528
Ser Gly Asn Pro Glu Asp Gly Gly Thr Gly Ala Gly Asn Gln Ile Ile	
165 170 175	

TCA	ACT	GCT	CAA	GTA	AAC	TTA	TCC	CAG	CCG	GAT	GCA	GCT	ACA	CTT	AAA	576
Ser	Thr	Ala	Gln	Val	Asn	Leu	Ser	Gln	Pro	Asp	Ala	Ala	Thr	Leu	Lys	
180					185					190					195	
GTC	GAT	GGA	GTA	TCT	GAT	CAT	AAA	TCT	GTC	TTT	GAT	GGC	GGA	GAC	GGT	624
Val	Asp	Gly	Val	Ser	Asp	His	Lys	Ser	Val	Phe	Asp	Gly	Gly	Asp	Gly	
				200					205					210		
ACA	GTT	TAT	CAA	AAT	ATT	CAG	CAA	TTT	ATC	GAT	GAA	GGC	AAG	TGG	ATT	672
Thr	Val	Tyr	Gln	Asn	Ile	Gln	Gln	Phe	Ile	Asp	Glu	Gly	Lys	Trp	Ile	
			215					220					225			
TCA	GGT	GAT	AAC	CAT	ACT	TTA	AGA	GAC	CCT	CAC	TAT	GTT	GAA	GAT	AAG	720
Ser	Gly	Asp	Asn	His	Thr	Leu	Arg	Asp	Pro	His	Tyr	Val	Glu	Asp	Lys	
		230					235					240				
GGC	CAT	AAA	TAT	CTT	GTC	TTT	GAA	GCG	AAT	ACT	GGA	ACA	ACA	GAT	GGT	768
Gly	His	Lys	Tyr	Leu	Val	Phe	Glu	Ala	Asn	Thr	Gly	Thr	Thr	Asp	Gly	
	245					250				255						
TAT	CAA	GGC	GAT	CAG	TCT	TTC	AAT	AAT	AAA	GCT	TAC	TAT	GGC	GGA	AGT	816
Tyr	Gln	Gly	Asp	Gln	Ser	Phe	Asn	Asn	Lys	Ala	Tyr	Tyr	Gly	Gly	Ser	
260					265					270					275	
GAC	GTC	TTC	TTC	CAG	AAT	GAA	AAA	AAT	AAA	CTG	CTT	CAA	AGT	CCT	AAA	864
Asp	Val	Phe	Phe	Gln	Asn	Glu	Lys	Asn	Lys	Leu	Leu	Gln	Ser	Pro	Lys	
				280					285					290		
AAA	CAA	ATT	GCT	TCT	TTA	GCG	AAT	GGT	GCA	TTA	GGC	ATT	GTT	GAA	TTG	912
Lys	Gln	Ile	Ala	Ser	Leu	Ala	Asn	Gly	Ala	Leu	Gly	Ile	Val	Glu	Leu	
			295					300					305			
GCC	GAT	GAC	TAT	ACA	GTG	AAA	AGT	GTT	ATG	AAA	CCA	TTA	GTC	GCA	TCA	960
Ala	Asp	Asp	Tyr	Thr	Val	Lys	Ser	Val	Met	Lys	Pro	Leu	Val	Ala	Ser	
		310					315					320				
AAC	ACA	GTA	GCA	GAT	GAA	GTC	GAA	CGC	GCC	AAT	ATA	TTT	AAA	ATG	AAT	1008
Asn	Thr	Val	Ala	Asp	Glu	Val	Glu	Arg	Ala	Asn	Ile	Phe	Lys	Met	Asn	
		325				330					335					
AAT	AAA	TGG	TAT	CTA	TTC	ACG	GAT	TCA	AGA	GGA	TCC	AAA	ATG	ACG	AGT	1056
Asn	Lys	Trp	Tyr	Leu	Phe	Thr	Asp	Ser	Arg	Gly	Ser	Lys	Met	Thr	Ser	
340					345					350					355	
GAT	GGA	ATT	AAC	GAC	AAA	GAT	GTT	TAT	ATG	CTA	GGG	CCC	GGA	GGC	GAC	1104
Asp	Gly	Ile	Asn	Asp	Lys	Asp	Val	Tyr	Met	Leu	Gly	Pro	Gly	Gly	Asp	
				360					365					370		
TCC	TTA	AAT	GGC	CCA	CAC	AAC	CCG	ATA	AAT	GAA	ACT	GGA	CTT	GTA	TTG	1152
Ser	Leu	Asn	Gly	Pro	His	Asn	Pro	Ile	Asn	Glu	Thr	Gly	Leu	Val	Leu	
			375					380					385			
AAC	ATG	AAT	CTT	GAC	CCT	GCT	GAT	CTC	ACA	CAC	ACT	TAC	TCT	CAT	TGC	1200
Asn	Met	Asn	Leu	Asp	Pro	Ala	Asp	Leu	Thr	His	Thr	Tyr	Ser	His	Cys	
		390					395					400				
GGT	ATC	CCG	CAC	CCT	GAA	GGT	AAT	AAT	GTG	GTA	CTC	ACA	AGT	TAT	ATG	1248
Gly	Ile	Pro	His	Pro	Glu	Gly	Asn	Asn	Val	Val	Leu	Thr	Ser	Tyr	Met	
	405					410					415					
ACG	AAT	AGA	GGC	TTC	TAT	CCA	GAA	CAT	CAC	TCT	CAC	CTG	CGG	GAC	AAG	1296
Thr	Asn	Arg	Gly	Phe	Tyr	Pro	Glu	His	His	Ser	His	Leu	Arg	Asp	Lys	
					425					430					435	
CTT	GGG	GTT	AAT	ATT	AAA	GGG	TCT	GAC	ACA	TCT	GGA	GGA	GAA	AAT	AGT	1344
Leu	Gly	Val	Asn	Ile	Lys	Gly	Ser	Asp	Thr	Ser	Gly	Gly	Glu	Asn	Ser	
				440					445					450		

TCC GGA CAA GGA CAA TTC CCA  
 Ser Gly Gln Gly Gln Phe Pro  
 455

1365

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 2408 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: genomic DNA
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Bacillus sp.
  - (B) INDIVIDUAL ISOLATE: V230 (FERM BP-5054)
- (ix) FEATURE:
  - (A) NAME/KEY: 5'UTR
  - (B) LOCATION: 1..360
  - (C) IDENTIFICATION METHOD: E
  - (A) NAME/KEY: signal peptide
  - (B) LOCATION: 361..456
  - (C) IDENTIFICATION METHOD: S
  - (A) NAME/KEY: mat peptide
  - (B) LOCATION: 457..1821
  - (C) IDENTIFICATION METHOD: S
  - (A) NAME/KEY: 3'UTR
  - (B) LOCATION: 1822..2408
  - (C) IDENTIFICATION METHOD: E

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CGGGGAAAT	ACTAGATTCC	AATTGGCCAG	ACTTCCAGT	TGGTGTAAGA	GAAGAGTTCG	60
GACTGCCAAT	GCAGCTGTGC	GTAAGAAAAC	AGCTTACTCA	TGAGCAATTA	CTAGAAGAAT	120
TTCAAAAGTC	CTGGGATAAG	GCCAAGTCCA	CTTTGAAATA	AACTTTTCAG	CCTCTGTGTG	180
GGGGCTTTTT	TGTTTTTATT	TATTTCAACT	GCAAGTGGTC	CATCCCCTAT	ATCAATTTAA	240
GACGAAATTC	TAATCAATCC	ATGCCATCCC	CAATAAACTC	GTCTCCTCT	ATACTTTTAA	300
TTAATAAGAA	ACTATCAAGA	GCTTTCTTAT	CAAATTCATA	CATATCCAAG	GAGGGAGACG	360
ATG AAC TTC AAA AGA TTG GCG AAA AAA GCA GCT GCC GTA ACC TTC AGG						408
Met Asn Phe Lys Arg Leu Ala Lys Lys Ala Ala Ala Val Thr Phe Arg						
-30		-25		-20		
ACT GCT ATA TTA GTA GGA GCG GAC GGA CCG CAT ATT TTT GCG CAG CAA						456
Thr Ala Ile Leu Val Gly Ala Asp Gly Pro His Ile Phe Ala Gln Gln						
-15		-10		-5		
ATG AAC AGC GGG GAC TAC AAG GAA GAC TAT GGT TTT GCC CAT ATT ACA						504
Met Asn Ser Gly Asp Tyr Lys Glu Asp Tyr Gly Phe Ala His Ile Thr						
1		5		10		15
CGC GCT GAC ATG CTA AAA ATT CCA GGA CAA CAA AAC AGT CCT CAA TTT						552
Arg Ala Asp Met Leu Lys Ile Pro Gly Gln Gln Asn Ser Pro Gln Phe						
20		25		30		
AAA GTG CCT CAA TTC AAT GCA TCA GCA ATC AAA AAC ATT GAT TCG GCA						600
Lys Val Pro Gln Phe Asn Ala Ser Ala Ile Lys Asn Ile Asp Ser Ala						
35		40		45		
AAA GGG TAT GAT AAG TCA GGC AAC TTA ATA GAT TTA GAT GTA TGG GAT						648
Lys Gly Tyr Asp Lys Ser Gly Asn Leu Ile Asp Leu Asp Val Trp Asp						
50		55		60		

AGC Ser 65	TGG Trp	CCA Pro	CTG Leu	CAA Gln	AAC Asn 70	GCT Ala	GAT Asp	GGT Gly	ACT Thr	GCG Ala 75	GCA Ala	AAT Asn	TAT Tyr	CAT His	GGA Gly 80	696
TAT Tyr	CAC His	ATC Ile	GTC Val	TCC Ser 85	GCT Ala	TTA Leu	GCA Ala	GGT Gly	GAC Asp 90	CCA Pro	AAA Lys	AAC Asn	AGT Ser	GAT Asp 95	GAT Asp	744
ACT Thr	CCA Pro	CTT Leu	CAT His 100	TTA Leu	TTC Phe	TAT Tyr	CAA Gln	AAA Lys 105	GTC Val	GGT Gly	GAT Asp	ACA Thr	TCG Ser 110	ATT Ile	GAC Asp	792
AGC Ser	TGG Trp	AAA Lys 115	AAT Asn	GCT Ala	GGA Gly	AGA Arg	GTA Val 120	TTT Phe	GAA Glu	GAT Asp	ATG Met	GAT Asp 125	AAA Lys	TTT Phe	GTT Val	840
CCA Pro	AAT Asn 130	GAT Asp	CCG Pro	TAT Tyr	CTT Leu	AAA Lys 135	TAT Tyr	CAA Gln	ACA Thr	CAG Gln	GAG Glu 140	TGG Trp	TCA Ser	GGT Gly	TCT Ser	888
GCT Ala 145	ACT Thr	TTA Leu	ACC Thr	AAA Lys	GAT Asp 150	GGC Gly	CAA Gln	GTC Val	CGT Arg	TTA Leu 155	TTC Phe	TAT Tyr	ACA Thr	GAT Asp	TAC Tyr 160	936
TCA Ser	GGT Gly	AAT Asn	CCT Pro	GAA Glu 165	GAT Asp	GGT Gly	GGA Gly	ACC Thr	GGT Gly 170	GCT Ala	GGT Gly	AAC Asn	CAA Gln	ATC Ile 175	ATT Ile	984
TCA Ser	ACT Thr	GCT Ala	CAA Gln 180	GTA Val	AAC Asn	TTA Leu	TCC Ser	CAG Gln 185	CCG Pro	GAT Asp	GCA Ala	GCT Ala	ACA Thr 190	CTT Leu	AAA Lys	1032
GTC Val	GAT Asp	GGA Gly 195	GTA Val	TCT Ser	GAT Asp	CAT His	AAA Lys 200	TCT Ser	GTC Val	TTT Phe	GAT Asp	GGC Gly 205	GGA Gly	GAC Asp	GGT Gly	1080
ACA Thr	GTT Val 210	TAT Tyr	CAA Gln	AAT Asn	ATT Ile	CAG Gln 215	CAA Gln	TTT Phe	ATC Ile	GAT Asp	GAA Glu 220	GGC Gly	AAG Lys	TGG Trp	ATT Ile	1128
TCA Ser 225	GGT Gly	GAT Asp	AAC Asn	CAT His	ACT Thr 230	TTA Leu	AGA Arg	GAC Asp	CCT Pro	CAC His 235	TAT Tyr	GTT Val	GAA Glu	GAT Asp	AAG Lys 240	1176
GGC Gly	CAT His	AAA Lys	TAT Tyr	CTT Leu 245	GTC Val	TTT Phe	GAA Glu	GCG Ala	AAT Asn 250	ACT Thr	GGA Gly	ACA Thr	ACA Thr	GAT Asp 255	GGT Gly	1224
TAT Tyr	CAA Gln	GGC Gly	GAT Asp 260	CAG Gln	TCT Ser	TTT Phe	AAT Asn	AAT Asn 265	AAA Lys	GCT Ala	TAC Tyr	TAT Tyr	GGC Gly 270	GGA Gly	AGT Ser	1272
GAC Asp	GTC Val	TTC Phe 275	TTC Phe	CAG Gln	AAT Asn	GAA Glu	AAA Lys 280	AAT Asn	AAA Lys	CTG Leu	CTT Leu	CAA Gln 285	AGT Ser	CCT Pro	AAA Lys	1320
AAA Lys	CAA Gln 290	ATT Ile	GCT Ala	TCT Ser	TTA Leu	GCG Ala 295	AAT Asn	GGT Gly	GCA Ala	TTA Leu 300	GGC Gly	ATT Ile	GTT Val	GAA Glu	TTG Leu	1368
GCC Ala 305	GAT Asp	GAC Asp	TAT Tyr	ACA Thr 310	GTG Val	AAA Lys	AGT Ser	GTT Val	ATG Met	AAA Lys 315	CCA Pro	TTA Leu	GTC Val	GCA Ala	TCA Ser 320	1416
AAC Asn	ACA Thr	GTA Val	GCA Ala	GAT Asp 325	GAA Glu	GTC Val	GAA Glu	CGC Arg	GCC Ala 330	AAT Asn	ATA Ile	TTT Phe	AAA Lys 335	ATG Met	AAT Asn	1464

AAT AAA TGG TAT CTA TTC ACG GAT TCA AGA GGA TCC AAA ATG ACG AGT	1512
Asn Lys Trp Tyr Leu Phe Thr Asp Ser Arg Gly Ser Lys Met Thr Ser	
340 345 350	
GAT GGA ATT AAC GAC AAA GAT GTT TAT ATG CTA GGG CCC GGA GGC GAC	1560
Asp Gly Ile Asn Asp Lys Asp Val Tyr Met Leu Gly Pro Gly Gly Asp	
355 360 365	
TCC TTA AAT GGC CCA CAC AAC CCG ATA AAT GAA ACT GGA CTT GTA TTG	1608
Ser Leu Asn Gly Pro His Asn Pro Ile Asn Glu Thr Gly Leu Val Leu	
370 375 380	
AAC ATG AAT CTT GAC CCT GCT GAT CTC ACA CAC ACT TAC TCT CAT TGC	1656
Asn Met Asn Leu Asp Pro Ala Asp Leu Thr His Thr Tyr Ser His Cys	
385 390 395 400	
GGT ATC CCG CAC CCT GAA GGT AAT AAT GTG GTA CTC ACA AGT TAT ATG	1704
Gly Ile Pro His Pro Glu Gly Asn Asn Val Val Leu Thr Ser Tyr Met	
405 410 415	
ACG AAT AGA GGC TTC TAT CCA GAA CAT CAC TCT CAC CTG CGG GAC AAG	1752
Thr Asn Arg Gly Phe Tyr Pro Glu His His Ser His Leu Arg Asp Lys	
420 425 430	
CTT GGG GTT AAT ATT AAA GGG TCT GAC ACA TCT GGA GGA GAA AAT AGT	1800
Leu Gly Val Asn Ile Lys Gly Ser Asp Thr Ser Gly Gly Glu Asn Ser	
435 440 445	
TCC GGA CAA GGA CAA TTC CCA TA GCGATTATCT CCCAATAAAA AGAAATGTCA	1853
Ser Gly Gln Gly Gln Phe Pro	
450 455	
CTGGCAAGAT CCTACCTTTT CCCCACCTT TTTTTTTAAA ATAAAGGGTT TTGTACCACC	1913
TTTAGAAGAA AAAAGAATCC TTGGCCCGGG CCAATTACCC ATTGCCGAAG GAGCTACCCA	1973
ATAAAAAAGA AATTTTCCCT TTTTAGGGGG GGAGTCCTT TTTTCTATCT TGGGGTGGGG	2033
ATTGTTGGCC CCCACCAGGG GACCTTTATT AATTTATAGC ATGAATCTGG CGATTTTGCC	2093
TGGCCTACTT ATTATCCAAG CCGCCCAGCC AAAAATATTA AACCAAGGAA TCCTAAGAGT	2153
AGGCGGTACA GTTTACCAAC ACCTGCTGGA GGGGATGGAT ATCGCCCTCA ATACCACTTT	2213
TCTGTTCCCG ACAAATGGAA AAATGATCCG CAGAGGCCCA TCTTTTTTGG GGGAAGTATC	2273
ATTACTATTA CCTTTACAAC AAAGACTATC CAGATGGAAA TGGTACGGAA TGGCGGCATG	2333
CAACGTCCGA AGATTTATTG CATTGGACGG ACGAAGGGAT TGCCATCCCG AAGTATACCA	2393
ATAAAAATGG TGATC	2408

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 26 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GAYTAYAARG ARGAYTAYGG NTTYGC

26